REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
09-03-2016	Conference Proceeding	-

4. TITLE AND SUBTITLE

A Label Propagation Approach for Detecting Buried Objectsin Handheld GPR Data

5a. CONTRACT NUMBER
W911NF-14-1-0589

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHORS
Graham Reid, Hichem Frigui

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of Louisville
2301 S. Third Street
Jouett Hall
Louisville, KY 40208 -1838

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
66411-CS.5

12. DISTRIBUTION AVAILIBILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

Detection of buried landmines and other explosive objects using ground penetrating radar (GPR) has been investigated for almost two decades and several classifiers have been developed. Most of these methods are based on the supervised learning paradigm where labeled target and clutter signatures are needed to train a classifier to discriminate between the two classes. Typically, large and diverse labeled training samples are needed to improve the performance of the classifier by overcoming noise and adding robustness and generalization to unseen examples. Unfortunately, even though unlabeled GPR data may be abundant, labeled data are often available in

15. SUBJECT TERMS

Buried Object Detection, Ground Penetrating Radar, Label Propagation

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Hichem Frigui
UU	UU	UU	UU		19b. TELEPHONE NUMBER 502-852-2009

Report Title

A Label Propagation Approach for Detecting Buried Objectsin Handheld GPR Data

ABSTRACT

Detection of buried landmines and other explosive objects using ground penetrating radar (GPR) has been investigated for almost two decades and several classifiers have been developed. Most of these methods are based on the supervised learning paradigm where labeled target and clutter signatures are needed to train a classifier to discriminate between the two classes. Typically, large and diverse labeled training samples are needed to improve the performance of the classifier by overcoming noise and adding robustness and generalization to unseen examples. Unfortunately, even though unlabeled GPR data may be abundant, labeled data are often available in small quantities as the labeling process is tedious and can be ambiguous for most of the data.

In this paper, we propose an algorithm for detecting landmines and buried objects that uses unlabeled data to help labeled data in the classification process. Our algorithm is graph-based and propagates the nodes' labels to neighboring nodes according to their proximity in the feature space. For labeled data, we use a set of prototypes that are extracted from a small set of labeled training samples. For unlabeled data, we use a collection of signatures that are extracted from the vicinity of the alarm being tested. This choice is based on the assumption that many spatially close signatures are expected to have similar features and thus, unlabeled samples can create dense regions that link different regions of the labeled samples and propagate their labels to test samples. In other words, unlabeled samples are explored to create a context for each test alarm.

To validate the proposed label propagation based classifier, we use it to detect buried explosive objects in GPR data collected by an experimental hand held demonstrator. We show that our approach is robust and computationally efficient to be used for both target discrimination and prescreening.

Conference Name: Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XXI, SPIE Defense and Secu

Conference Date: April 17, 2016

A Label Propagation Approach for Detecting Buried Objects in Handheld GPR Data

Graham Reid and Hichem Frigui

Multimedia Research Lab, CECS Dept.
University of Louisville, Louisville, KY 40292, USA;

ABSTRACT

Detection of buried landmines and other explosive objects using ground penetrating radar (GPR) has been investigated for almost two decades and several classifiers have been developed. Most of these methods are based on the supervised learning paradigm where labeled target and clutter signatures are needed to train a classifier to discriminate between the two classes. Typically, large and diverse labeled training samples are needed to improve the performance of the classifier by overcoming noise and adding robustness and generalization to unseen examples. Unfortunately, even though unlabeled GPR data may be abundant, labeled data are often available in small quantities as the labeling process is tedious and can be ambiguous for most of the data. In this paper, we propose an algorithm for detecting landmines and buried objects that uses unlabeled data to help labeled data in the classification process. Our algorithm is graph-based and propagates the nodes labels to neighboring nodes according to their proximity in the feature space. For labeled data, we use a set of prototypes that are extracted from a small set of labeled training samples. For unlabeled data, we use a collection of signatures that are extracted from the vicinity of the alarm being tested. This choice is based on the assumption that many spatially close signatures are expected to have similar features and thus, unlabeled samples can create dense regions that link different regions of the labeled samples and propagate their labels to test samples. In other words, unlabeled samples are explored to create a context for each test alarm. To validate the proposed label propagation based classifier, we use it to detect buried explosive objects in GPR data collected by an experimental hand held demonstrator. We show that our approach is robust and computationally efficient to be used for both target discrimination and prescreening.

Keywords: Buried Object Detection, Ground Penetrating Radar, Label Propagation

1. INTRODUCTION

Buried explosive objects detection and removal is important to preserving both military and civilian lives. While planting these objects is easy, detecting and removing them is relatively difficult. To combat this threat, many devices have been employed utilizing Ground Penetrating Radar (GPR),[1–5] electromagnetic induction (EMI),[6–8] and other remote sensing techniques.[9–11] Both vehicle-mounted and hand-held GPR sensing devices are currently employed for this purpose.

In general, buried explosive objects detection techniques using GPR data follow a two-step process: First, an anomaly detector (or pre-screener) is employed to process the data stream and identify regions of interest that correspond to locations with anomalous signatures. Second, a classifier (or an ensemble of classifiers) is used to assign a confidence value denoting whether or not the anomaly reflects the presence of buried target or of a clutter object. For real-time detection systems, pre-screening algorithms should be computationally very efficient and should be designed to achieve high probability of detection even at the expense of a false alarm rate higher than the total system requirements. False alarms detected by pre-screening algorithms are typically due to (sub-) surface non-target anomalies, sudden changes in antennae's height, or changes in the subsurface structure. The classifier involves a feature extraction step that aims to characterize the salient properties of each region of interest followed by a supervised learning algorithm that is trained to discriminate between the salient features of targets and non-targets. Classifiers can afford to be computationally more expensive since they process a much smaller subset of the data.

Further author information: Send correspondence to Hichem Frigui (h.frigui@louisville.edu)

Landmine detection techniques involving pre-screening and classification for GPR have evolved considerably over the last couple of decades. Some recent examples of pre-screeners that have been explored include least mean squares fitting,[12] ground alignment, and median filtering. In some cases, if the classifier can be made efficient enough, it can be used as a pre-screener.[13]

Several landmine detection techniques operate on GPR using edge-based features, essentially transforming landmine detection into an image analysis/computer vision problem. Two commonly applied features are Edge Histogram Descriptors (EHD)[14] and Histogram of Oriented Gradients HOG.[13] These features are simple yet capable of accurately representing the typical hyperbolic shape of buried objects in GPR. The EHD has been implemented in fielded systems mounted on a vehicle and has proven to be robust in detecting targets of different types and under various environmental conditions. HOG features represent edges at a finer details and capture a wider range of edge directions. However, this richer description can potentially reduce the generalization capabilities.

After features are extracted, the data are passed to a supervised classifier to assign confidences. Examples of feature-based discriminative classifiers that have been used in this application include K-Nearest Neighbor (KNN),[14] Hidden Markov Models,[15] Support Vector Machines,[16] Random Forest[13] and Multiple-Instance Learning.[17–19] Most of these techniques rely on constructing a predictive model to characterize data, and as a consequence are prone to over-fitting the training data.

To avoid over-fitting and reduce the computational complexity of the KNN classifier, fuzzy labeling can be used. Instead of using all of the training data, fuzzy/possibilistic KNN (p-KNN) summarizes the labeled data with few representative prototypes. Then, each prototype is assigned two independent soft label: one in the class of targets and another in the class of clutter. In addition, the distance of each nearest prototype is converted to a possibilistic membership values to weigh its contribution to the final confidence value.

In this paper, we argue that the EHD classifier could be improved by considering two additional factors. First, instead of treating the target and non-target prototypes independently from each other, we take into account their pairwise similarities. We construct a fully connected graph where the nodes are all the prototypes and the edges correspond to their similarities. Second, instead of treating the problem as a supervised learning one, we use additional unlabeled contextual data to help the labeled data in classifying test data. The proposed algorithm could be used as a classifier that operates only at locations of interest flagged by a prescreener. Alternatively, due to its computational efficiency, it could be used as a prescreener. In this paper, we treat it as a prescreener.

The rest of this paper is organizes as follows. In Section 2, we describe the data and our label propagation approach. In Section 3, we describe our experimental setup and compare our approach to the standard KNN and possibilistic KNN. For the label propagation, we analyze the results as we vary the number of unlabeled samples. Finally, Section 4 contains the conclusions.

2. BURIED OBJECT DETECTION USING LABEL PROPAGATION

2.1 GPR Data and Feature Extraction

The input data consists of a stream of raw GPR signatures collected by an experimental hand held demonstrator. This system collects 3 channels of data. Scans are collected as the operator moves the sensor down-track. For a fixed channel, the sequence at each down-track position contains 256 time samples (which are approximately related to depth) at which the GPR signal return is reported. The down-track scans are uniformly sampled to have one scan per 1 cm. The collected input data can be represented by a 3-Dimensional matrix of sample values, $S(z, x, y), z = 1, \cdots, 416; x = 1, \cdots, 3; y = 1, \cdots, N_{dt}$, where N_{dt} is the total number of collected scans, and the indicesz, x,and y represent depth, channel, and down-track positions respectively. In this paper, we only use data from channel 1

Figure 1 illustrates sample data consisting of 200 scans extracted from channel 1.

The high energy (dark band) around time sample 50 corresponds to the response from the air/ground interface. Time samples above this band correspond to the sensor signal propagating through air, and time samples below the sample correspond to propagation in the soil. The samples in Fig. 1 cover a region that includes one target (visible between time samples 75 and 125 and at down-track sample 118).

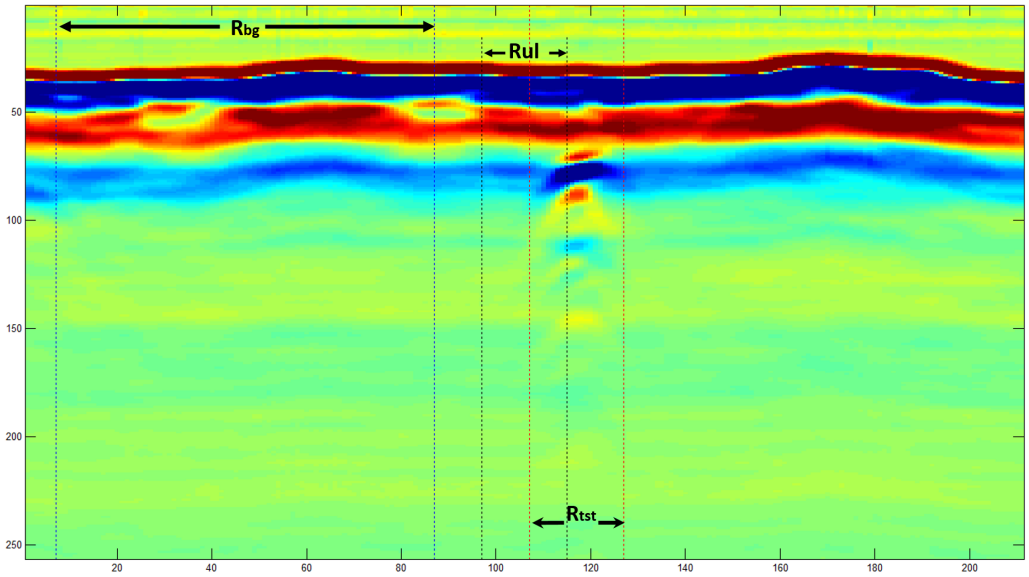


Figure 1. A sample of 210 GPR scans collected by an experimental hand held demonstrator that includes a target around scan 120 and visible between time samples 75 and 125. R_{tst} refers to the region being tested, R_{bg} refers to the region used to estimate the background statistics, and R_{ul} refers to the region used to extract additional unlabeled samples when testing R_{tst}.

First, we preprocess the data to remove stationary effects by subtracting the background. The background is dynamically estimated using a collection of scans prior to the location being tested. In Fig. 1, region R_{Bg} refers to the collection of scans used to estimate the background when region R_{tst} is being tested. Fig. 2 displays the test region from Fig. 1 (R_{tst}) after background subtraction.

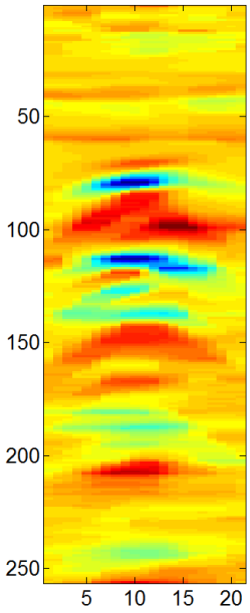


Figure 2. Test region R_{tst} from Fig. 1 after background subtraction.

Next, each region R_{tst} is represented by a feature vector that captures its salient features. Any of the existing features outlined earlier could be used. In this paper, we use a slightly modified EHD features.[14] First, no edges

are extracted from the cross-track direction (since the hand held demonstrator does not have a cross-track resolution comparable to the vehicle mounted GPR). Second, non-edge information is not explicitly included. Thus, each GPR test region is mapped to a 16-dimensional feature vector.

2.2 Prototypes Learning

We assume that we have a training data collection where each alarm is labeled as target or non-target/clutter using the ground truth. As in the EHD,[14] to reduce the size of the training samples and reduce the risk of over-fitting, we identify few representative prototypes that can summarize and capture the within-class variations. In particular, the target and clutter signatures are clustered separately, and only one representative alarm is selected from each cluster. Let $\mathbf{x}_i$, $i = 1, \cdots N_L$ denote the feature vector of the i^{th} prototypes where N_L is the total number of prototypes. Each prototype, $\mathbf{x}_i$, is assigned a fuzzy membership in the class of targets, y_i^t, and a fuzzy membership in the class of clutter y_i^c based on the distribution of the clusters.[14] Let $\mathbf{y}_i$=(y_i^t, y_i^c) denote the soft labels of prototype i.

2.3 Label Propagation

Let $(\mathbf{x}_1, \mathbf{y}_1), \cdots, (\mathbf{x}_{N_L}, \mathbf{y}_{N_L})$ be the N_L labeled prototypes identified from the labeled data, where $Y_L = \{\mathbf{y}_1, \cdots, \mathbf{y}_{N_L}\}$ are the class soft labels. Let $(\mathbf{x}_{N_L+1}, \mathbf{y}_{N_L+1}), \cdots, (\mathbf{x}_{N_L+N_U}, \mathbf{y}_{N_L+N_U})$ be the N_U unlabeled data samples where $Y_U = \{\mathbf{y}_{N_L+1}, \cdots, \mathbf{y}_{N_L+N_U}\}$ are unobserved. In our application, the unlabeled data correspond to features extracted from the test region and its neighborhood as will be explained later. The problem is to estimate Y_U from X_L, X_U, and Y_L.

First, we construct a fully connected graph where each $\mathbf{x}_i \in \{X_L \cup X_U\}$ is a node. The edge connecting two nodes i and j, w_{ij} reflects the similarity between the nodes and is computed using

$$w_{ij} = \exp\Big(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma^2}\Big) \tag{1}$$

The main idea behind the label propagation algorithm is to let the soft labels of a node to propagate to all nodes through the edges.[20] A larger edge weight will make the propagation easier. Nodes that belong to the labeled subset will keep their original labels while nodes from the unlabeled samples will use the learned (i.e. propagated) labels.

Let $\mathbf{T} = [\mathbf{T}_{ij}]$ be a $(N_L + N_U) \times (N_L + N_U)$ probabilistic transition matrix defined as

$$\mathbf{T}_{ij} = P(j \to i) = \frac{w_{ij}}{\sum_{k=1}^{N_L+N_U} w_{kj}}. \tag{2}$$

In (2), T_{ij} is the probability to jump from node j to node i. The row-normalized transition matrix, $\bar{\mathbf{T}} = [\bar{\mathbf{T}}_{ij}]$, is obtained using

$$\bar{\mathbf{T}}_{ij} = \frac{\mathbf{T_{ij}}}{\sum_{k=1}^{N_L+N_U} \mathbf{T}_{ik}} \tag{3}$$

Next, we split $\bar{\mathbf{T}}$ into 4 sub-matrices along the labeled/unlabeled nodes boundary, i.e.,

$$\bar{\mathbf{T}} = \begin{bmatrix} \bar{\mathbf{T}}_{LL} & \bar{\mathbf{T}}_{LU} \\ \bar{\mathbf{T}}_{UL} & \bar{\mathbf{T}}_{UU} \end{bmatrix} \tag{4}$$

It can be shown[20] that $\mathbf{Y_U}$ can be obtained using

$$\mathbf{Y_U} = \big(\mathbf{1} - \bar{\mathbf{T}}_{\mathbf{UU}}\big)^{-\mathbf{1}} \bar{\mathbf{T}}_{\mathbf{UL}} \mathbf{Y_L} \tag{5}$$

Thus, given a training data set with partially labeled samples, labels for the unlabeled subset can be learned in a direct and non-iterative way.

2.4 Selection of Unlabeled Data

As it can be seen in Fig. 1, the target signature does not span all depth values. Typically, depending on the mine type and burial depth, the mine signature may extend over 25 to 80 depth values. That is, it may cover no more than 10% of the total depths under consideration. Thus, extracting one global feature vector from the alarm may not discriminate between targets and clutter signatures effectively. To avoid this limitation, each scan location is tested at multiple depth values. We slide a (21scans $\times$60depths) window size along the depth axis with a 66% overlap between 2 consecutive signatures. A total of 10 signatures are extracted for each sample. A similar testing approach has been used for target detection algorithms developed for vehicle-mounted GPR systems. In these systems, the multiple windows are either tested independently of each other and then their partial confidence values are combined,[14] or are combined into a bag representation and a multiple instance learning algorithm is used.[17,19,21,22] In our proposed approach, at each location, we treat the features extracted from the 10 windows as unlabeled data and solve for their labels simultaneously using (5). To obtain the final confidence value, we sort the 10 partial confidence values and average the largest three.

Typically, to improve the generalization of the learning classifier and improve its computational efficiencies, only few representative prototypes are selected. These prototypes tend to be sparse in the high-dimensional feature space. Thus, additional unlabeled data that can potentially create denser regions between the 10 test samples and some of the prototypes can improve the accuracy of the classifier. In this paper, we extract an additional 30 unlabeled samples extracted from a region that precedes the scan being tested (region R_{Ul} in Fig. 1).

2.5 Parameter Setting

In addition to selecting unlabeled data samples, the label propagation algorithm requires the specification of σ (used in equation (1)). If σ is too small, the labels assigned to the unlabeled samples will be influenced by only the nearest labeled sample. Similarly, if σ is too large, all of the labeled samples will contribute almost equally to the labeling of all of $\mathbf{X_U}$. As a result, the class probabilities assigned to each unlabeled sample will be simply the class average of the labeled samples. In this paper, we use the same heuristic used in.[20] First, we find a minimum spanning tree over all data points using the Euclidean distances d_{ij}. Then, we find the smallest tree edge that connects two nodes with different labels. If d^0 is the length of this edge, then we let $\sigma = d^0/3$.

3. EXPERIMENTAL SETUP AND RESULTS

The proposed label propagation prescreener was applied to analyze data of buried landmines and other explosive objects collected using a Ground Penetrating Radar (GPR) sensor. The data was collected using an experimental hand held demonstrator from outdoor test lanes. Two collections were available. One was used exclusively for training while the second one was used for testing. The training set includes data collected from 5 different lanes spanning 336.4 square meters and consisting of 58 targets. First, we identify image regions in the (depth, down-track) plane that correspond to target locations and that have strong (visible) GPR signatures. These signatures form the labeled training target data. This data is then summarized, using self-organizing map (SOM),[23] by 20 representative target prototypes. Next, we randomly selected many non-target image regions that do not overlap with the true locations of the targets. These make the labeled non-target training data. This data is also summarized by 20 non-target (i.e., background or clutter) prototypes. Each of the 40 prototypes is assigned a fuzzy label ($\in [0,1]$) in the class of targets and another soft label in the class of non-targets. These labels are based on the content of each cluster and its distance from other target/non-target prototypes.

The second data collection, used for testing only, consists of 12 lanes spanning 441.7 square meters and consisting of 174 targets. Each lane is tested at every other scan. At each location, we test sub-signatures extracted from 10 overlapping windows (21 scans $\times$ 60 depths) covering all 256 depth values. The average confidence of the top 3 confidence windows is taken as the final confidence value.

To illustrate the advantages of treating the prototypes as a fully connected graph, we compare the results of the proposed label propagation to those obtained using the K-NN and the possibilistic K-NN.[14] All algorithms use the same set of prototypes. The standard K-NN uses crip labels (1 for targets and 0 for non-targets), while the fuzzy K-NN and label propagation algorithms use soft labels. To illustrate the advantage of including

additional unlabeled samples, we test 3 versions of the Label Propagation (LP). The first one, LP1, treats all 10 test windows independently from each other and tests only one window at a time. In other words, LP1 solves equation (5) for only one unlabeled samples. The second version, LP10, tests all 10 windows simultaneously, i.e., LP10 solves (5) for 10 unlabeled samples. The third version, LP with unlabeled samples (or simply LP), tests all 10 windows simulateneously in addition to 30 unlabeled samples extracted from a region just prior to the scan being tested (region R_{UL} in Figure 1).

The scoring is performed in terms of probability of detection (PD) versus false alarms rate (FAR). Confidence values are thresholded at different levels to produce the receiver operating characteristics (ROC) curve. The results are shown in Figure 3. First, we note that at low PD/FAR, the PD of most algorithms are comparable (these are the relatively stronger targets that are easy to detect). However, at higher PD, the variation in PD can reach 10%. The crisp K-NN has the worst performance. The fuzzy K-NN outperforms the crisp KNN because it uses soft labels and it takes into account the distance of the nearest neighbors (through a possibilistic membership function). As for the 3 LP versions, LP1 did not perform as well as the possibilistic KNN. The benefits of using a fully connected graph are not sufficient to match the benefits of using a possibilistic membership function in the KNN. Lp10, on the other hand, outperforms the possibilistic KNN and LP with additional unlabeled samples has the best overall performance.

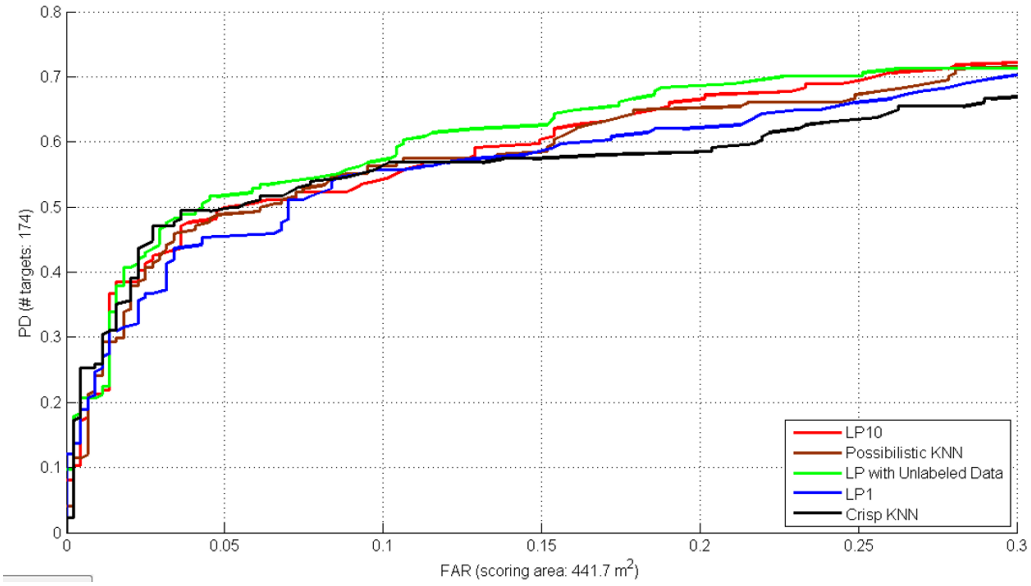


Figure 3. Comparison of the performance of the K-NN, Possibilistic K-NN, and Label Propagation algorithms. For the Label propagation, we report the results when only one window is tested at a time (LP1), all 10 windows tested simultaneously (LP10), and when additional unlabeled samples are included .

To gain more insights, for each algorithm we identify the threshold that leads to a FAR= 0.15. Then, we identify and analyze targets/non-targets that were detected/missed by each algorithm. Figure 4 shows a sample target that was missed by the KNN and detected by LP1 and LP10. In this figure, the partial confidence values of each algorithm, at each of the 10 depths, are displayed as a bar plot. For all 3 algorithms the confidence is higher around the middle depths where the target signature is strong. The vertical dashed line in each bar plot indicates the confidence threshold at which FAR=0.15. For the middle depth bin, LP1 assigns a confidence much larger than the KNN. This is because, unlike KNN, LP1 considers the pairwise similarities among all target/non-target prototypes. Another important observation is that LP10 generates high confidence values for 3 (not 1 as in LP1) depth bins. The reason is that when scoring all 10 sub-signatures simultaneously, some of their features will create "bridge" points to some target prototypes. Thus, more of these prototypes will contribute to their confidence values.

Figure 5 shows a sample false alarm detected by the KNN that is barely missed by LP1 and completely missed by LP10. In the latter case, the features of the 10 sub-signatures, when tested simultaneously, create

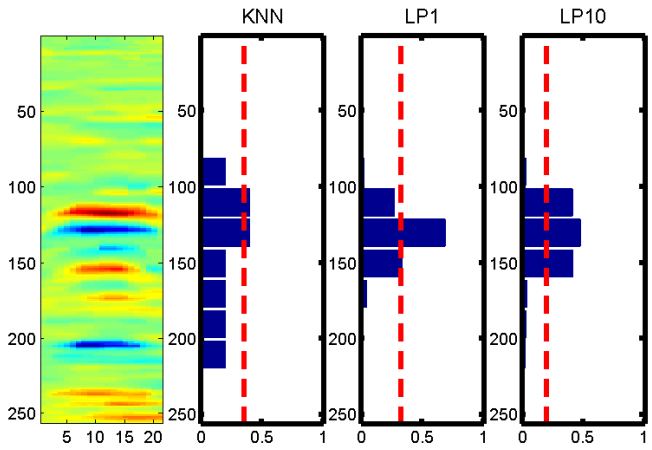


Figure 4. Sample target signature that was missed by the KNN (at FAR=0.15) and detected by LP1 and LP10.

"bridge" points to more non-target prototypes. Thus, their confidence values are reduced significantly.

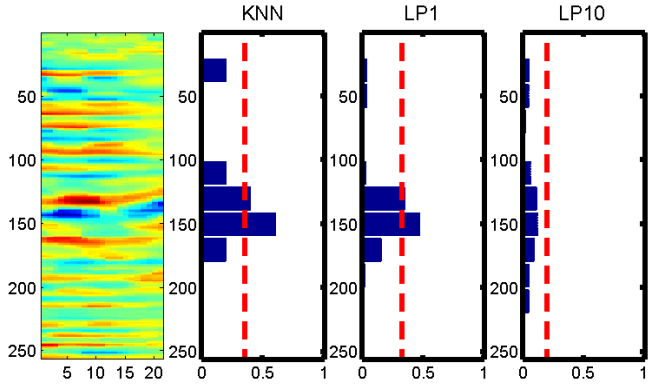
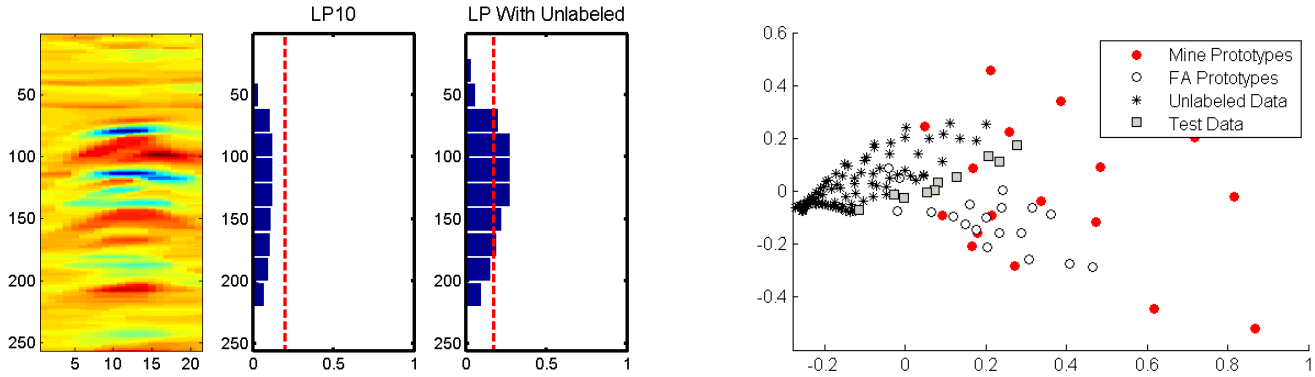


Figure 5. Sample non-target signature (FA) that was detected by the KNN (at FAR=0.15), barely missed by LP1, and has a much lower confidence using LP10.

In Figure 6, we illustrate the advantage of using additional unlabeled samples (extracted from a region that precedes the test scan) as a context. In Figure 6(a), we display a sample target signature that was missed by LP10 and detected by LP with additional unlabeled points. In Figure 6(b), we display a 2-D scatter of all prototypes, the 10 test samples (at different depths), and the additional 50 unlabeled samples. Here, the original 16-dimensional features are projected into 2-D using multi-dimensional scaling. As it can be seen, many of the unlabeled samples create denser regions that bridge the test points to target prototypes. This explains the increased confidence values at many depth bins. Figure 7 shows the opposite behavior for a non-target. Here, the unlabeled samples create denser regions that bridge test points to non-target prototypes. As a result, the confidence values at most depths has decreased significantly. In this case, this alarm is detected as a false alarm (at FAR=0.15) by LP10 and not by LP with unlabeled points.

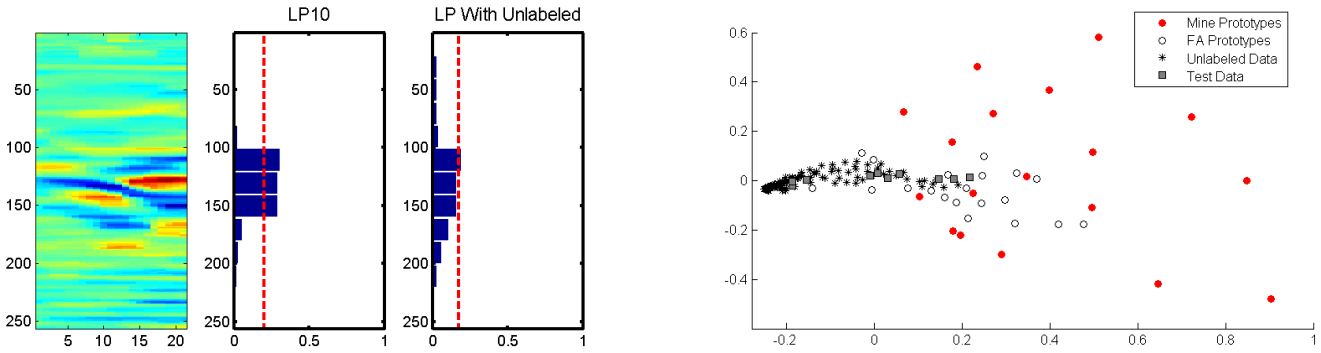
4. CONCLUSIONS

In this paper, we have presented a semi-supervised algorithm to detect buried landmines and other explosive objects using GPR data. This algorithm, called label propagation, relies on both labeled and unlabeled data. It is based on the assumption that data points that are close to each other in the feature space tend to have similar class labels. For labeled data, we use target and non-target prototypes that are extracted from a small set of labeled training samples. For unlabeled data, we use a collection of signatures that are extracted from the vicinity



(a) (b)

Figure 6. Illustration of the advantage of using additional unlabeled samples in detecting targets. (a) a sample target and the confidence values assigned at each of the 10 depths by LP10 and LP with more unlabeled data. (b) Scatter plot of the target/non-target prototypes, test samples at the 10 depth bins, and additional 30 unlabeled samples. Here the 16 dimensional feature space is projected into 2-D using multi-dimensional scaling.



(a) (b)

Figure 7. Illustration of the advantage of using additional unlabeled samples in reducing false alarms. (a) a sample target and the confidence values assigned at each of the 10 depths by LP10 and LP with more unlabeled data. (b) Scatter plot of the target/non-target prototypes, test samples at the 10 depth bins, and additional 30 unlabeled samples. Here the 16 dimensional feature space is projected into 2-D using multi-dimensional scaling.

of the alarm being tested. Each labeled and unlabeled sample is treated as a node in a fully connected graph. The edge between two nodes reflects their similarity. The nodes labels' are then propagated to neighboring nodes according to their proximity in the feature space.

To validate the proposed approach, we use it to detect buried explosive objects in GPR data collected by an experimental hand held demonstrator. Simple edge histogram descriptors were used as features to construct the graph. This resulted in a detection algorithm that is computationally very efficient and can be used as a prescreener that processes all scans of the GPR data stream. We showed that additional unlabeled samples, used as a context for test samples, can improve the performance of the detector. For targets, unlabeled data can create dense regions that bridge the test samples to more target prototypes, thus increasing the probability of detection. Similarly, for non-targets, unlabeled data can create dense regions that bridge the test samples to more clutter/background prototypes, thus reducing the false alarm rate.

Currently, only simple edge histogram features have been used to construct the edges of the graph. As future work, we will investigate other features such as Histogram of Oriented Gradients HOG.[13] Other future work will include optimizing the selection of σ (used in equation (1)). We will also investigate learning one σ for each dimension in order to identify relevant features.

Acknowledgment

This work was supported in part by U.S. Army Research Office Grants Number W911NF-13-1-0066 and W911NF-14-1-0589. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office, or the U.S. Government.

REFERENCES

[1] J. Farrell, T. C. Havens, K. Ho, J. M. Keller, T. T. Ton, D. C. Wong, and M. Soumekh, "Evaluation and improvement of spectral features for the detection of buried explosive hazards using forward-looking ground-penetrating radar," in *SPIE Defense, Security, and Sensing*, pp. 83571C–83571C, International Society for Optics and Photonics, 2012.

[2] L. E. Besaw and P. J. Stimac, "Deep learning algorithms for detecting explosive hazards in ground penetrating radar data," in *SPIE Defense+ Security*, pp. 90720Y–90720Y, International Society for Optics and Photonics, 2014.

[3] N. P. Singh and M. J. Nene, "Buried object detection and analysis of gpr images: Using neural network and curve fitting," in *Emerging Research Areas and 2013 International Conference on Microelectronics, Communications and Renewable Energy (AICERA/ICMiCR), 2013 Annual International Conference on*, pp. 1–6, IEEE, 2013.

[4] M. Ndoye and J. M. Anderson, "An mm-based algorithm for l 1-regularized least squares estimation in gpr image reconstruction," in *Radar Conference (RADAR), 2013 IEEE*, pp. 1–6, IEEE, 2013.

[5] M. Loewer, N. Wagner, and J. Igel, "Prediction of gpr performance in soils using broadband dielectric spectroscopy," in *Near Surface Geoscience 2013-19th EAGE European Meeting of Environmental and Engineering Geophysics*, 2013.

[6] L. Carin, H. Yu, Y. Dalichaouch, A. R. Perry, P. V. Czipott, and C. E. Baum, "On the wideband emi response of a rotationally symmetric permeable and conducting target," *Geoscience and Remote Sensing, IEEE Transactions on* **39**(6), pp. 1206–1213, 2001.

[7] P. Torrione and L. Collins, "Performance comparison of automated induction-based algorithms for landmine detection in a blind field test," *Subsurface Sensing Technologies and Applications* **5**(3), pp. 121–150, 2004.

[8] A. Pinar, M. Masarik, T. C. Havens, J. Burns, B. Thelen, and J. Becker, "Approach to explosive hazard detection using sensor fusion and multiple kernel learning with downward-looking gpr and emi sensor data," in *SPIE Defense+ Security*, pp. 94540B–94540B, International Society for Optics and Photonics, 2015.

[9] E. V. McDonald and R. Schumer, "Use of ground-based lidar for detection of ied command wires on typical desert surfaces," in *Military Geosciences and Desert Warfare*, pp. 297–309, Springer, 2016.

[10] H. Petersson, D. Gustafsson, J. Hedborg, and D. Letalick, "Field trials for development of multiband signal processing solutions to detect disturbed soil," in *SPIE Security+ Defence*, pp. 964909–964909, International Society for Optics and Photonics, 2015.

[11] M. P. Nelson, A. Basta, R. Patil, O. Klueva, and P. J. Treado, "Development of a handheld widefield hyperspectral imaging (hsi) sensor for standoff detection of explosive, chemical, and narcotic residues," in *SPIE Defense, Security, and Sensing*, pp. 872605–872605, International Society for Optics and Photonics, 2013.

[12] P. A. Torrione, C. S. Throckmorton, and L. M. Collins, "Performance of an adaptive feature-based processor for a wideband ground penetrating radar system." IEEE Trans. Aerospace and Electronic Systems (in press).

[13] P. Torrione, K. Morton, R. Sakaguchi, and L. Collins, "Histograms of oriented gradients for landmine detection in ground-penetrating radar data," *Geoscience and Remote Sensing, IEEE Transactions on* **52**, pp. 1539–1550, March 2014.

[14] H. Frigui and P. Gader, "Detection and discrimination of land mines in ground-penetrating radar based on edge histogram descriptors and a possibilistic k-nearest neighbor classifier," *Trans. Fuz Sys.* **17**, pp. 185–199, Feb. 2009.

[15] P. Gader, M. Mystkowski, and Y. Zhao, "Landmine detection with ground penetrating radar using hidden markov models," *IEEE Trans. Geoscience and Remote Sensing* **39**, pp. 1231–1244, 2001.

[16] A. Hamdi, O. Missaoui, and H. Frigui, "An svm classifier with hmm-based kernel for landmine detection using ground penetrating radar," in *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*, pp. 4196–4199, July 2010.

[17] A. Manandhar, P. Torrione, L. Collins, and K. Morton, "Multiple-instance hidden markov model for gpr-based landmine detection," *Geoscience and Remote Sensing, IEEE Transactions on* **53**, pp. 1737–1745, April 2015.

[18] A. Karem and H. Frigui, "A multiple instance learning approach for landmine detection using ground penetrating radar," in *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*, pp. 878–881, IEEE, 2011.

[19] S. Yuksel, J. Bolton, and P. Gader, "Multiple-instance hidden markov models with applications to landmine detection," *Geoscience and Remote Sensing, IEEE Transactions on* **53**, pp. 6766–6775, Dec 2015.

[20] X. Zhu and Z. Ghahramani, "Learning from labeled and unlabeled data with label propagation," tech. rep., 2002.

[21] A. B. Khalifa and H. Frigui, "Multiple instance fuzzy inference," in *Fuzzy Systems (FUZZ-IEEE), 2015 IEEE International Conference on*, (2015).

[22] A. Khalifa and H. Frigui, "A multiple instance neuro-fuzzy inference system for fusion of multiple landmine detection algorithms," in *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International*, pp. 4312–4315, July 2015.

[23] T. Kohonen, *Self-Organization and Associative Memory*, Springer Verlag, 1989.